

**THE EFFECTS OF VERB- AND EVENT-BASED KNOWLEDGE ON DIRECT OBJECT  
PREDICTION**

by

**Evelyn Arko Milburn**

B.A. in Psychology, University of California, Davis, 2010

Submitted to the Graduate Faculty of  
the Kenneth P. Dietrich School of Arts  
and Sciences in partial fulfillment of  
the requirements for the degree of  
Master of Science in Psychology

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH  
DIETRICH SCHOOL OF ARTS  
AND SCIENCES

This thesis was presented

by

Evelyn Arko Milburn

It was defended on

April 30, 2014

and approved by

Charles Perfetti, PhD, University Professor, Psychology

Michael Walsh Dickey, PhD, Assistant Professor, Communication Science and Disorders

Committee Chair: Tessa Warren, PhD, Associate Professor, Psychology

Copyright © by Evelyn Arko Milburn

2014

# **THE EFFECTS OF VERB- AND EVENT-BASED KNOWLEDGE ON DIRECT OBJECT PREDICTION**

Evelyn Arko Milburn, M.S.

University of Pittsburgh, 2014

Previous research suggests that comprehenders rapidly predict upcoming words on the basis of both language and event knowledge (e.g., Kamide et al, 2003), but we know little about how these kinds of knowledge interact in driving prediction. We report a visual-world eye-tracking study comparing prediction driven by the combination of event knowledge and a weakly constraining verb to prediction driven by event knowledge and a verb that places strong semantic constraints on its arguments. College-aged adults (n=36) viewed photographs of natural scenes while listening to sentences. Participants anticipated upcoming direct objects similarly regardless of the presence of a strongly constraining verb. We discuss these findings in light of literature on selectional restriction violations (Paczynski & Kuperberg, 2012; Warren & McConnell, 2007) and research suggesting that event-related knowledge is used very early in sentence comprehension (McRae & Matsuki, 2009), as well as the relationship between prediction and production.

## TABLE OF CONTENTS

<b>PREFACE.....</b>	<b>IX</b>
<b>1.0 INTRODUCTION.....</b>	<b>1</b>
<b>1.1 CURRENT STUDY .....</b>	<b>5</b>
<b>2.0 EXPERIMENT 1.....</b>	<b>8</b>
<b>2.1 METHODS.....</b>	<b>8</b>
<b>2.1.1 Stimuli.....</b>	<b>8</b>
<b>2.1.2 Norming.....</b>	<b>10</b>
<b>2.2 PROCEDURE .....</b>	<b>14</b>
<b>2.3 RESULTS .....</b>	<b>15</b>
<b>2.4 DISCUSSION.....</b>	<b>16</b>
<b>3.0 EXPERIMENT 2.....</b>	<b>17</b>
<b>3.1 METHODS.....</b>	<b>17</b>
<b>3.1.1 Stimuli.....</b>	<b>17</b>
<b>3.1.2 Norming.....</b>	<b>17</b>
<b>3.2 PROCEDURE .....</b>	<b>19</b>
<b>3.3 RESULTS .....</b>	<b>20</b>
<b>3.3.1 Verb Bin.....</b>	<b>21</b>
<b>3.3.2 Noun Bin.....</b>	<b>21</b>

3.3.3	First Post-Noun Bin .....	22
3.3.4	Second Post-Noun Bin .....	22
3.3.5	Verbal Fluency .....	22
3.3.6	Backwards Digit Span .....	23
4.0	DISCUSSION .....	24
BIBLIOGRAPHY .....		29

## LIST OF TABLES

Table 1: Experiment 1 Cloze Norm .....	11
Table 2: Experiment 1 Picture Norm .....	12
Table 3: Experiment 1 Event Norm .....	13
Table 4: Experiment 1 Mean Verb Segment and Determiner Lengths .....	14
Table 5: Experiment 2 Cloze Norm .....	19
Table 6: Experiment 2 Picture Norm .....	19

## LIST OF FIGURES

Figure 1: Image Accompanying Event Constrained and Event Control Sentences.....	9
Figure 2: Image Accompanying Verb Constrained and Verb Control Sentences .....	10
Figure 3: Mean Fixation Proportions For Each Condition During Each Time Window.....	20
Figure 4: Mean Latency to Fixate Target .....	23



## **PREFACE**

I wish to thank Katherine Martin and Alba Tuninetti for help with visual stimuli creation, Joni Keating and Zac Ekves for help running participants, and members of the LABlab for help recording and measuring audio stimuli. I thank my academic advisor, Tessa Warren, for her help in developing this project and for comments on previous versions of this manuscript. I also thank my Master's committee for their insightful comments.

## 1.0 INTRODUCTION

Examining prediction of upcoming words during online sentence processing can provide important insight into what sources of information comprehenders draw on when understanding language. These different sources of information may, in turn, make dissociable contributions to predictive processing. Comprehenders make predictions about likely upcoming words based on the following: verb-related information (Altmann & Kamide, 1999; Staub, Abbott, & Bogartz, 2012), world knowledge about an agent's likely actions (Kamide, Altmann, & Haywood, 2003), the sentential context of the predicted word, (Federmeier & Kutas, 2005; Federmeier, 2007), and priming by semantic association (Kukona, Fang, Aicher, Chen, & Magnuson, 2011). Findings indicate that comprehenders' predictions are influenced by information ranging from the lexical level to the contextual level; however, we are just beginning to unravel how these sources of information compete and conspire to drive prediction. The purpose of the present research is to investigate whether there exists linguistic—specifically, verb-based—knowledge that is used in a dissociable way from knowledge about likely upcoming events. To illuminate this possible dissociation, we explore the potential different contributions of verb- and event-based knowledge to direct object prediction.

Sources disagree regarding whether or not there exists specifically linguistic knowledge that can drive language comprehension separately from world knowledge. Jackendoff (2002) argues that there is no easily-drawn line between lexical information and world knowledge.

Under this view, all language input is processed with reference to a store of generalized world knowledge—including information about events and their standard participants—which is gained by personal experience (Ferretti, McRae, & Hatherell, 2001; McRae, Hare, Elman, & Ferretti, 2005). Results from several studies support this perspective; for example, verbs prime typical agents, patients, and instruments (Ferretti et al, 2001), which can, in turn, prime verbs (McRae et al, 2005). Additionally, the combination of a verb with an aspect indicating that the event being described is ongoing (“*was skating*”) primes typical locations associated with that event (Ferretti, Kutas, & McRae, 2007). According to McRae and colleagues, language comprehension is based on the rapid combination of world knowledge constraints imposed by cues from multiple words (McRae & Matsuki, 2009), and there is no specifically linguistic information that can influence language comprehension in a manner dissociable from world knowledge.

Katz & Fodor (1963), however, proposed that all words carry with them a set of semantic features, which are accessed immediately during language comprehension. Some verbs constrain their set of possible arguments to ones with particular semantic features; these constraints are called selectional restrictions. Under this view, world knowledge, which is stored separately from lexical semantic features, would be used along a different time-course. Although there is now significant evidence that world knowledge can be used immediately during sentence comprehension (McRae & Matsuki, 2009), several studies investigating the effects of selectional restriction violations (SRVs) have found hints of separable effects of selectional restrictions during language comprehension. For example, sentences containing SRVs elicit more disruption to eye movements than sentences containing equally severe violations of world knowledge (Warren, & McConnell, 2007; Warren, Milburn, Patson, & Dickey, under review), even when

the SRVs are embedded in a rich supportive context—for example, a loaf of bread learning math after Harry Potter casts a spell to animate it (Warren, McConnell, & Rayner, 2008). Results from studies like these, as well as studies examining verb processing in people with aphasia (Nakano & Blumstein, 2004; Myers & Blumstein, 2005), suggest that verb-based information might be used differently from world knowledge during language comprehension.

A number of studies have investigated how different kinds of information guide prediction, finding that participants can make predictions based on verbs, combinatory effects of a sentence's agent and verb, and global context information drawn from properties of the visual scene itself. Using the Visual World Paradigm (VWP; an eyetracking paradigm able to provide sensitive time-course information about language processing) Altmann & Kamide (1999) found that the minimal semantic requirements imposed by a verb on its potential direct objects can guide eye movements to the appropriate object in the visual scene, even before that object is mentioned. However, global information about the likely progression of a scene can also be used to drive prediction independently from local verb information. Again using the VWP, Kamide et al (2003) found predictive eye movements to targets that satisfied the semantic requirements of the verb and were appropriate given the agent and action described in the linguistic stimulus. The observed results could not be explained solely by the semantic constraints of the verb or by knowledge about the agent's likely actions, but instead by the combinatory effect of the verb's requirements given the accompanying scene. Finally, Spivey and colleagues demonstrated that the visual context of an utterance can help resolve syntactic ambiguities (Spivey, Tanenhaus, Eberhard, & Sedivy, 2002). This suggests that language comprehension can be influenced by properties of an accompanying array of objects, consistent with a view in which world

knowledge plays a large role in language comprehension. Taken together, the results of these studies show that prediction can be driven by multiple sources.

Predictive eye movements may be influenced by global context information, like the scene and world knowledge about likely events, but several studies also reveal small effects of individual lexical items, especially verbs, on prediction even when presented concurrently with conflicting contextual information. Kukona et al (2011) found increased looks to a thematically appropriate role filler even when that role had already been filled in the linguistic input—for example, looks to a picture of a policeman increased after “arrested” in “Toby arrested the crook”. Similarly, Borovsky, Elman, and Fernald (2012) found that participants made anticipatory fixations on images that were associated with the agent or the verb individually, even if those images were inconsistent with the constraints imposed by the agent and verb in combination.

These local effects of individual lexical items indicate that world knowledge is not the only kind of information being used to drive prediction. In the studies detailed above, world knowledge is activated by the combination of all the words in the linguistic stimulus. Under a view in which predictive eye movements are driven solely by world knowledge, words in combination should not prime words that are inconsistent with that combination: “The pirate chases the \_\_\_\_” should not prime “cat” because, although “cat” is a plausible patient of “chase,” cats are typically not chased by pirates. The fact that multiple studies have shown these verb-specific local effects on prediction despite a conflicting global context is congruent with a potential dissociative use of verb- and event-based knowledge when making predictions.

## 1.1 CURRENT STUDY

Previous research has demonstrated that participants are able to integrate information from multiple sources very quickly in order to make predictions about upcoming words in a sentence (Altmann & Kamide, 1999; Kamide et al, 2003; Borovsky et al, 2012; Kukona et al, 2011), but further study is needed to determine whether different sources of information make separable contributions to prediction, and, if so, what those contributions might be. The current study aims to investigate whether event knowledge and verb knowledge differentially drive prediction by examining eye movements in the visual world. Previous research (e.g. Warren & McConnell, 2007) suggests that violations of constraining verb information are detected earlier than violations of world knowledge. One hypothesis that could explain these early effects of verb constraint on comprehension is that verb constraints generate earlier or stronger predictions than world knowledge. The current study tests this hypothesis by investigating whether verb constraints provide a boost to prediction over that provided by world knowledge. We compare conditions containing the same amount of event information but which vary in the strength of the constraining verb.

Each experimental stimulus provides the participant with different kinds of information with which to make predictions. In the event-constrained experimental condition we pair unconstrained verbs with pictures depicting a highly-predictive event—for example, a picture of a bride at a wedding paired with the sentence “Someone will fling the \_\_\_\_\_” together predict the target word “flowers,” verified by a cloze norm. Because of the lack of semantic information provided by each sentence’s agent and verb, the linguistic stimuli are non-predictive of the target when presented without the corresponding image. Prediction-guiding information is contained in the combination of the scene and the low-constraint verb. In the verb-constrained condition the

scenes are paired with verbs that place strong semantic constraints on their potential objects—for example, the sentence “Someone will strum the \_\_\_\_\_” predicts “guitar” as an object without the help of an image, again verified by a cloze norm. Each experimental stimulus is paired with a control stimulus in which the only characteristic changed is the verb. Previous studies have used combinations of verb, agent, and scene information (Kamide et al, 2003; Borovsky, 2012) to activate world knowledge; the present design allows us to examine any potential additional boost to prediction given by constraining verb information over information provided by the scene.

Several studies have identified an influence of verb information on direct object prediction (Altmann & Kamide, 1999; Kamide et al, 2003; Staub et al, 2012), but thus far the effects of verb information have not been fully isolated: verb information can combine with information about the agent (Kukona et al, 2011; Borovsky et al, 2012) and the agent’s likely actions (Kamide et al, 2003) to drive eye movements to a particular object in the visual scene. In order to avoid these combinatory effects, in the present study we use sentences containing a semantically empty agent (“someone”) rather than the semantically-rich subjects used in previous work, thus removing any potential influence of semantic information from words other than the verb.

A potential concern with the VWP is that characteristics of the images used may artificially drive eye movements and confound any observed results: in most studies, the visual input is so sparse that the participant may increase their reliance on the linguistic input when making predictions. To address this issue, we use photographs instead of clip art figures (cf. Staub et al, 2012), thereby imbuing our visual stimuli with rich event-based information.

We report the results of two experiments with the same design and most of the same stimuli. As detailed below, the results of Experiment 1 were potentially influenced by confounds contained within the visual stimuli; we therefore performed Experiment 2 to clarify our results.



## **2.0 EXPERIMENT 1**

### **2.1 METHODS**

#### **2.1.1 Stimuli**

Critical stimuli consisted of 32 naturalistic scenes, each accompanied by a constraining sentence and a control sentence (see Figures 1a and 1b for examples). Half of the constraining sentences contained verbs that limited their potential direct objects to a single object in the scene (Figure 1b). The other half of the constraining sentences contained verbs with very few semantic requirements of their direct objects; in this condition, the constraint was due to the combination of the verb with the event depicted in the scene (Figure 1a). All control sentences contained a verb that could plausibly refer to many objects in the visual scene and was not predictive of a specific object given the agents and objects in the scene. All sentences consisted of a semantically empty subject (“someone”), a future-tense transitive verb, and a direct object.

Nine of the visual stimuli were obtained from Staub et al (2012), and the remaining 23 visual stimuli were either drawn from flickr’s pool of Creative Commons-licensed images (<https://www.flickr.com/creativecommons/>) or staged by the investigators. We also created 64 filler stimuli, each of which consisted of a picture paired with a single sentence. Although all fillers used semantically empty subjects, they varied both in post-verbal structure and in their

relationship to the accompanying scene. Filler images were either obtained from Staub et al (2012) or were Creative Commons-licensed images from flickr. All images in the experiment were resized to 1024x768 pixels.



**Figure 1: Image Accompanying Event Constrained and Event Control Sentences**

Event-Constrained sentence: “Someone will fling the flowers.”

Event-Control sentence: “Someone will love the flowers.”



**Figure 2: Image Accompanying Verb Constrained and Verb Control Sentences**

Verb-Constrained Sentence: “Someone will strum the guitar.”

Verb-Control Sentence: “Someone will hide the guitar.”

### **2.1.2 Norming**

All linguistic stimuli were normed in order to make sure that they were appropriately constrained or unconstrained. 18 students from the University of Pittsburgh student community participated for course credit. All were 18 years of age or older and were native speakers of English.

Participants filled out a questionnaire consisting of 64 sentences of the form “Someone will touch the \_\_\_\_” by completing each sentence with an appropriate word. The sentences were presented to the subjects in (pseudo-) random order, with no more than two items from the same condition occurring sequentially. We calculated the proportion of trials in which the target direct

object for each stimulus was provided. In some cases, we consolidated similar responses based on whether they had the same referent or were unlikely to be differentiated in the naturalistic scene. For example, for the item “Someone will erase the \_\_\_\_” we consolidated the answers “board,” “chalkboard,” and “blackboard.” In cases where answers were consolidated, proportions were calculated based on the consolidated scores. Table 1 shows idealized and actual mean target response proportions in the cloze norm. There was a significant main effect of sentence type ( $F(1, 12)=32.715$ ;  $p<.01$ ) and constraint ( $F(1, 12)=50.824$ ;  $p<.01$ ). There was also a significant interaction of sentence type and constraint ( $F(1, 12)=40.410$ ;  $p<.01$ ). Mean cloze scores were higher for the event constrained condition than for the event control condition ( $t(14)=2.329$ ;  $p<.05$ ). Mean cloze scores were likewise higher for the verb constrained condition than for the verb control condition ( $t(12)=7.207$ ;  $p<.05$ ). Additionally, mean cloze scores in the verb constrained condition were higher than in the event constrained condition ( $t(15)=6.552$ ;  $p<.05$ ). However, mean cloze scores for the two control conditions did not differ ( $t(12)=1.477$ ;  $p=.165$ ).

Table 1: Experiment 1 Cloze Norm: Idealized and Actual Mean Target Response Proportions and Standard Errors per Condition

**Table 1: Experiment 1 Cloze Norm**

Condition	Ideal Mean	Actual Mean	Standard Error
Event-Constrained	0	.067	.02
Event-Control	0	.007	.001
Verb-Constrained	1	.59	.08
Verb-Control	0	0	0

A second round of norms was completed using participants ( $n=28$ ) who had not completed the first round of norming. This round of norms was identical in procedure to the first round except that each sentence was accompanied by its corresponding naturalistic scene,

presented using a PowerPoint slideshow. Participants were told to complete each sentence based on the event they saw in the image. Proportions were calculated for these norms using the same procedure as in the first round of norms. Table 2 shows idealized and actual mean target responses in the picture norm. There was a significant main effect of constraint ( $F(1, 15)=231.51$ ;  $p<.01$ ), but no effect of verb type ( $p=.53$ ) or interaction ( $p=.57$ ). Mean scores in the event constrained condition were higher than in the event control condition ( $t(15)=9.66$ ;  $p<.05$ ). Likewise, scores in the verb constrained condition were higher than scores in the verb control condition ( $t(15)=12.59$ ;  $p<.05$ ). However, there was no significant difference in mean score in the two constrained conditions ( $t(15)=1.359$ ;  $p=.19$ ) or in the two control conditions ( $t(15)=.068$ ;  $p=.95$ ).

Table 2: Experiment 1 Picture Norm: Idealized and Actual Mean Target Response Proportions

**Table 2: Experiment 1 Picture Norm**

Condition	Ideal Mean	Actual Mean	Standard Error
Event-Constrained	1	.92	.04
Event-Control	0	.26	.05
Verb-Constrained	1	.98	.02
Verb-Control	0	.33	.07

A final round of norms was completed using participants ( $n=15$ ) who had not completed either of the first two rounds of norming. This norm was intended to measure how much non-linguistic semantic information (event information) was present in just the picture stimulus, with no cues from the accompanying linguistic stimulus. Participants viewed each image on a PowerPoint slideshow and completed sentences of the form “Someone will \_\_\_\_” based on what they thought would happen next in the image. Proportions were calculated based on how often

the participant provided the target, a reasonable synonym, or a semantically-related event when completing the sentence. Table 3 shows means and standard errors for the event norm. *T*-tests comparing mean target proportions showed that there was no significant difference in target response proportions between the event pictures and the verb pictures ( $t(15) = 1.611, p = .13$ ).

Table 3: Experiment 1 Event Norm: Mean Target Response Proportions and Standard Errors per Condition

**Table 3: Experiment 1 Event Norm**

Condition	Mean	Standard Error
Event	.44	.09
Verb	.25	.06

Since we used different verbs in the different conditions of our experiment, it was necessary to verify their frequencies. Raw verb frequencies were obtained from CELEX. There was a marginal effect of stimulus type on verb frequency ( $p=.06$ ). Verbs in the verb-constrained conditions ( $M = 861.25$ ) were less frequent than verbs in the event-constrained condition ( $M = 2024.88$ ) ( $t(30) = 2.166; p<.05$ ). This difference is a potential concern; since lower frequency verbs are slower to access, predictions driven by lower frequency verbs may take longer to appear in eye movements. However, it is exactly these lower frequency verbs that are more constraining, and should lead to stronger prediction.

All audio stimuli were recorded by a female native speaker of English. Table 4 shows mean verb segment and determiner lengths for each of the four conditions. One-way ANOVAs determined that no significant differences existed among the four groups for either verb segment length ( $F(3, 60) = 1.15$ ) or determiner length ( $F(3, 60) = 1.52$ ). Interest areas were created by drawing regions around each target object that extended approximately one degree of visual

angle from the target’s borders. Because our critical comparisons involved the verbs in the linguistic stimulus, there was only one target object per scene, and no competitors (cf. Staub et al, 2012).

Table 4: Experiment 1: Mean Verb Segment and Determiner Lengths

**Table 4: Experiment 1 Mean Verb Segment and Determiner Lengths**

Condition	Verb Segment	Standard Error	Determiner	Standard Error
Event-Constrained	580.00	24.63	251.25	10.20
Event-Control	565.63	19.50	253.75	12.11
Verb-Constrained	605.63	26.33	245.63	11.87
Verb-Control	592.97	21.36	278.12	12.46

## 2.2 PROCEDURE

Participants’ eyes were tracked using an Eyelink 1000 tracker (SR Research Ltd., Toronto, Ontario, Canada) with a sampling rate of 1 ms and a spatial resolution of less than a 30-min arc. Participants viewed stimuli binocularly on a monitor approximately 63 cm from their eyes. Head movements were minimized using forehead and chin rests. After explaining the format of the experiment to the participants, we calibrated the eye tracker using a 13-point fixation stimulus. In each trial, the visual stimulus was presented to the participant first, followed by the audio stimulus after a 1000 ms delay, thereby giving participants time to extract event-related information from the scene. A single-point drift correction was performed after every trial, as well as a full 13-point recalibration every 24 trials. Audio stimuli were presented to participants

via two speakers positioned at either side of the viewing monitor. The experiment lasted between 20 and 30 minutes.

We constructed two lists of stimuli. Each participant saw every visual stimulus, but heard only one of the two possible accompanying audio stimuli. Stimuli were presented to participants in random order.

Some theories of language production (e.g. Pickering & Garrod, 2007) propose a relationship between language production and predictive processing; this possible relationship is supported by studies of older adults, finding more young-like predictive behavior in older adults with higher language production skills (Federmeier, 2007; DeLong, Groppe, Urbach, & Kutas, 2012). To test theories connecting prediction and production, we administered a test of verbal fluency, which measures language production ability, to our participants. They first named as many animals as they could within a one-minute interval, then repeated the exercise with fruits; these are complementary measures of semantic category fluency. The total numbers of exemplars produced for each category were averaged to determine an overall verbal fluency score. In accordance with the results found for older adults, we expect that participants with higher verbal fluency will also show an advantage in prediction.

## **2.3 RESULTS**

We analyzed latency of first fixation to the target in a time window beginning at the onset of the critical verb and lasting until the end of the trial. If the participant was already fixating the target region during verb onset, the second fixation to the target was used (c.f. Staub et al, 2012), as this is the first fixation that could logically be driven by verb information. We removed trials



containing fixations longer than 2000 ms, as well as trials for which the latency of first fixation to the target was outside two standard deviations above the mean for that condition. These measures resulted in 14.13% of trials being discarded.

Because the nature of our visual stimuli (naturalistic scenes instead of clip-art images) does not allow inclusion of explicit distractor objects, our critical comparisons were between conditions rather than between target and distractor objects in the same scene (c.f. Staub et al, 2012). There was a significant main effect of verb type such that the mean latency to fixate the target was faster in the verb constrained and verb control conditions, but this effect was only significant by subjects ( $F_1(1, 39)=15.49$ ,  $p<.05$ ;  $F_2(1, 15)=1.124$ ,  $p=.17$ ). There was also a significant main effect of constraint such that mean latency to fixate the target was faster in the two constrained conditions, but this was only significant by items ( $F_1(1, 39)=3.508$ ,  $p=.07$ ;  $F_2(1, 15)=9.936$ ,  $p<.05$ ). There was no significant interaction ( $F_1(1, 39)=.010$ ,  $p=.92$ ;  $F_2(1, 15)=.015$ ,  $p=.90$ ).

## **2.4 DISCUSSION**

Although these data hint at an effect of overall constraint, whether based in the verb or in the depicted event, on prediction, the results are not strong enough to allow us to draw any definite conclusions. An investigation of individual items suggested that nine picture stimuli (five from the event-constrained condition, four from the verb-constrained condition) were visually confusing or were not clearly associated with an event. Therefore, we removed or replaced several picture stimuli and re-ran the experiment.

## **3.0 EXPERIMENT 2**

### **3.1 METHODS**

#### **3.1.1 Stimuli**

The nine poor visual stimuli from Experiment 1 were replaced with pictures staged by the experimenters. The accompanying audio stimuli remained the same, as did the remaining images. We also removed two items from analysis. Upon reflection, we determined that one of the verbs in the event-constrained condition (“serve ball”) was not sufficiently unconstrained in the sentence cloze norm, and the item was therefore removed. To balance our lists, we also removed the verb-constrained stimulus (“burp baby”) that had scored the lowest on the cloze and picture norms. All filler stimuli were unchanged.

#### **3.1.2 Norming**

Because we changed several visual stimuli, we renormed our images to ensure that participants were able to correctly choose the target offline. In a procedure identical to the previous experiment, participants (n=20) saw a PowerPoint slideshow of each image and completed sentences based on what they saw in the image. Cloze scores were calculated for these norms using the same procedure as in the first experiment. Table 6 shows idealized and actual

mean responses for the picture norm. There was a significant main effect of constraint such that participants provided the target object more frequently in the constrained conditions than in the control conditions ( $F(1,14) = 209.27$ ;  $p < .05$ ). There was no significant main effect of sentence type ( $F(1,14) = 2.52$ ;  $p = .14$ ) and no interaction ( $F(1,14) = .04$ ;  $p = .84$ ).

Table 5 shows idealized and actual mean responses for the sentence-only cloze norm, without the two discarded items. Reanalysis of the original cloze norms showed main effects of sentence type ( $F(1, 11) = 28.42$ ;  $p < .05$ ), and constraint ( $F(1, 11) = 44.65$ ;  $p < .05$ ), as well as a significant interaction ( $F(1, 11) = 35.38$ ;  $p < .05$ ). Paired-sample *t*-tests showed that participants provided the target object more frequently in the verb constrained condition more frequently than in the event constrained condition ( $t(14) = 6.21$ ;  $p < .05$ ), the verb control condition ( $t(11) = 6.75$ ;  $p < .05$ ), and the event control condition ( $t(13) = 7.49$ ;  $p < .05$ ).

We reanalyzed verb segment and determiner length after removing the two discarded items. Verb length did not vary reliably by sentence type ( $F(1, 14) = 4.53$ ;  $p = .052$ ), or constraint ( $F(1, 14) = .002$ ;  $p = .97$ ), and there was no interaction ( $F(1, 14) = .08$ ;  $p = .78$ ). Likewise, determiner length did not vary reliably by sentence type ( $F(1, 14) = .97$ ;  $p = .34$ ), or constraint ( $F(1, 14) = 2.76$ ;  $p = .12$ ), and there was no interaction ( $F(1, 14) = 2.94$ ;  $p = .11$ ).

Finally, we reanalyzed verb frequencies (obtained from CELEX). There were no significant effects of sentence type ( $F(1, 14) = .62$ ;  $p = .45$ ), or constraint ( $F(1, 14) = 2.77$ ;  $p = .12$ ), and no interaction ( $F(1, 14) = .01$ ;  $p = .91$ ).

Table 5: Experiment 2 Cloze Norm: Idealized and Actual Mean Target Response Frequencies and Standard Errors per Condition

**Table 5: Experiment 2 Cloze Norm**

Condition	Ideal Mean	Actual Mean	Standard Error
Event-Constrained	0	.08	.03
Event-Control	0	.01	.01
Verb-Constrained	1	.61	.09
Verb-Control	0	.00	.00

Table 6: Experiment 2 Picture Norm: Idealized and Actual Mean Target Response Frequencies and Standard Errors per Condition

**Table 6: Experiment 2 Picture Norm**

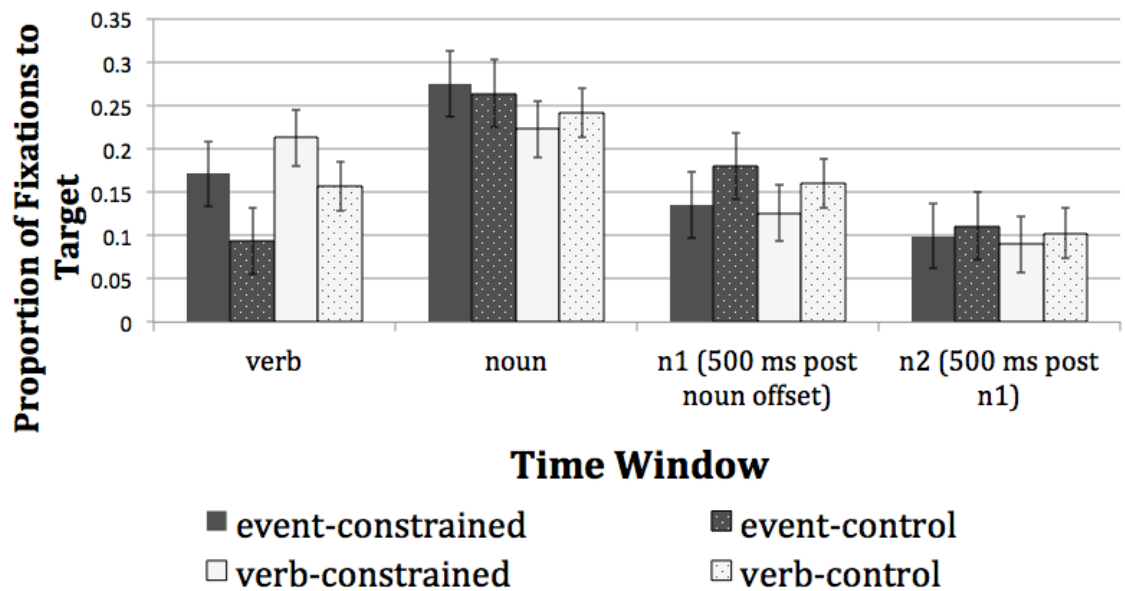
Condition	Ideal Mean	Actual Mean	Standard Error
Event-Constrained	1	.86	.03
Event-Control	0	.24	.06
Verb-Constrained	1	.95	.02
Verb-Control	0	.36	.09

### 3.2 PROCEDURE

36 undergraduate students from the University of Pittsburgh completed the experiment for course credit. None had participated in Experiment 1. The same procedure was used for Experiment 2 as for Experiment 1. As in Experiment 1, we administered a test of verbal fluency. For this experiment we added a test of backwards digit span, to have an individual difference measure that was not expected to be related to prediction.

### 3.3 RESULTS

We removed trials containing fixations longer than 2000 ms (12 trials), as well as trials in which the target was never fixated (93 trials). This resulted in 9.7% of trials being discarded; 975 total trials were analyzed. Proportions of fixations to the target were calculated for each of four time bins after the onset of the verb. Figure (2) shows mean proportion of fixations on the target during each bin.



**Figure 3: Mean Fixation Proportions For Each Condition During Each Time Window**

Figure 2: Mean fixation proportions for each condition during each time window. Comparisons marked with a star are significant at  $p < .05$ .

### **3.3.1 Verb Bin**

The verb bin began 200 ms after the onset of the verb and ended 200 ms after the onset of the noun; because eye movements take approximately 200 ms to plan and execute in responses to a stimulus, 200 ms post-verb onset was the earliest point at which fixations to the target could have been driven by verb information. There were significantly more looks to the target in the two constrained conditions than in the two control conditions in this bin ( $F_1(1, 35)=25.43, p<.05$ ;  $F_2(1, 14)=18.29, p<.05$ ). There were also significantly more looks to the target in the two verb conditions than in the two event conditions ( $F_1(1, 35)=10.96, p<.05$ ;  $F_2(1, 14)=8.65, p<.05$ ). However, there was no interaction of sentence type and constraint ( $F_1(1, 35)=.77, p=.39$ ;  $F_2(1, 14)=.120, p=.73$ ).

### **3.3.2 Noun Bin**

The noun bin began 200 ms after the onset of the noun and ended 200 ms after the offset of the noun. There was no effect of constraint ( $F_1(1, 35)=.08, p=.77$ ;  $F_2(1, 14)=1.09, p=.31$ ). There were significantly more looks to the target in the two event conditions than in the two verb conditions, but this effect was only significant by subjects ( $F_1(1, 35)=10.96, p<.05$ ;  $F_2(1, 14)=1.08, p=.32$ ). There was no interaction of sentence type and constraint ( $F_1(1, 35)=1.17, p=.29$ ;  $F_2(1, 14)=1.59, p=.23$ ).

### **3.3.3 First Post-Noun Bin**

The first post-noun bin began 200 ms after the offset of the noun and ended 700 ms after the offset of the noun. There were significantly more looks to the target in the two control conditions than in the two constrained conditions in this bin ( $F_1(1, 35)=8.55, p<.05$ ;  $F_2(1, 14)=9.89, p<.05$ ). There was no effect of sentence type ( $F_1(1, 35)=1.41, p=.24$ ;  $F_2(1, 14)=1.62, p=.22$ ), and no interaction of sentence type and constraint ( $F_1(1, 35)=.14, p=.71$ ;  $F_2(1, 14)=.11, p=.75$ ).

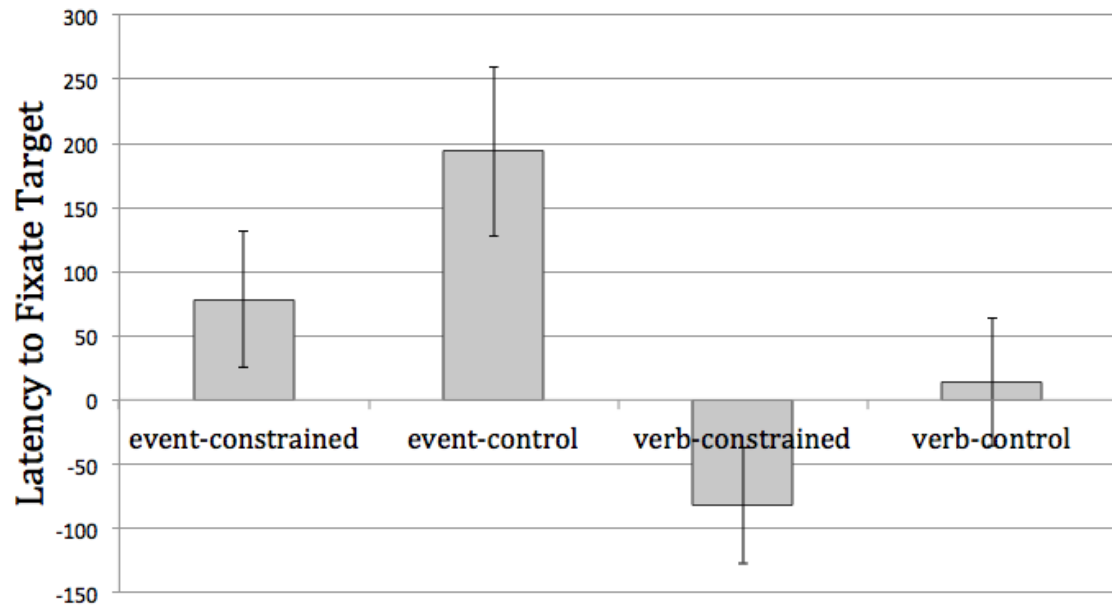
### **3.3.4 Second Post-Noun Bin**

The second post-noun bin began 700 ms after the offset of the noun and ended 1200 ms after the offset of the noun. There were no significant effects of constraint ( $F_1(1, 35)=1.86, p=.18$ ;  $F_2(1, 14)=2.97, p=.11$ ) or sentence type ( $F_1(1, 35)=.68, p=.42$ ;  $F_2(1, 14)=.001, p=.98$ ), and no significant interaction ( $F_1(1, 35)=.003, p=.95$ ;  $F_2(1, 14)=.001, p=.98$ ).

### **3.3.5 Verbal Fluency**

We observed a wide range of verbal fluency scores (11-26.5). Verbal fluency did not significantly predict time-locked noun onset latency ( $B = -10.62$ ;  $t(142) = 1.45$ ;  $R^2 = .015$ ;  $p=.15$ ). However, verbal fluency significantly predicted time to fixate the target after verb onset (time-locked verb latency) when all trials were entered into the equation ( $B = -17.08$ ;  $t(142) = 2.02$ ;  $R^2 = .028$ ;  $p<.05$ ). We also examined the effect of verbal fluency on amount of prediction by subtracting time-locked verb onset latency in the constrained conditions from time-locked verb onset latency in the control conditions. Verbal fluency did not significantly predict the

amount of prediction ( $B = 3.099$ ;  $t(34) = .268$ ;  $R^2 = .002$ ;  $p = .79$ ). Finally, we used verbal fluency to predict fixation proportions collapsed across constrained or control conditions in each time window (cf. Hintz, Meyer, & Huettig, 2014)). The only reliable effect was in the first post-noun bin for the constrained conditions; here verbal fluency was a significant predictor of fixation proportions ( $B = -15.434$ ;  $t(34) = 2.409$ ;  $R^2 = .146$ ;  $p < .05$ ).



**Figure 4: Mean Latency to Fixate Target**

Figure 3: Mean latency to fixate target, measured backwards from noun onset.

### 3.3.6 Backwards Digit Span

Backwards digit span did not significantly predict time-locked verb onset latency when all trials were entered into the equation ( $B = -4.85$ ;  $t(142) = 1.31$ ;  $R^2 = .012$ ;  $p = .193$ ).



## 4.0 DISCUSSION

We investigated the potential different contributions of verb- and event-based knowledge to direct object prediction by analyzing proportion of fixations to a target in the visual world. Research on impossibility detection during reading (Warren & McConnell, 2007; Warren et al, in prep) and semantic relatedness (Paczynski & Kuperberg, 2012), suggests that constraining verb information makes independent contributions to language comprehension and may play a role in prediction over and above that played by event knowledge. However, this study did not find any differences in prediction related to the presence of a constraining verb. Rather, participants predicted the direct object in both event-constrained and verb-constrained conditions, as measured by proportion of fixations to the target during the verb, suggesting that the primary driver of prediction in this experiment was event-based information gained from the visual scene. These results are consistent with evidence that event-based knowledge is used very early in language comprehension (McRae & Matsuki, 2009).

Although several previous studies have shown verb constraint effects on direct object prediction (Altmann & Kamide, 1999; Staub et al, 1012), they have not disentangled the individual effects of the verb's constraints, the sentence's agent, and the characteristics of the scene. Event and verb knowledge are tightly intertwined—after all, verbs describe events—and difficult to fully separate. This tight connection between event and verb knowledge means that the effects of constraining verb information on direct object prediction found in previous studies

are actually effects of a combination of verb and event knowledge. Both Altmann & Kamide (1999) and Staub et al (2012) used stimuli containing event information that could have been used to drive prediction along with the effects of their constraining verbs, and thus event and verb knowledge are confounded in their studies. For example, both studies used sentences containing semantically meaningful agents: although agents like “the boy” and “the woman” may seem relatively neutral, initial norms in the present study found that participants greatly changed their cloze responses based on even these seemingly unconstrained agents (“someone” holds cups, doors, bags, phones, etc., but “the woman” most frequently holds babies or purses). Neither Altmann & Kamide (1999) nor Staub et al (2012) manipulated effects of sentential agent, but these semantically-meaningful agents, along with the images, could have provided enough event information to drive prediction (cf. Kamide et al, 2003). Characteristics of the visual stimuli used in previous studies could also have influenced the apparent effects of verb information on prediction. Although Staub and colleagues used visual stimuli containing rich non-linguistic semantic information, similar to the present study, they did not include a manipulation attempting to disentangle verb information from event information, and thus the scenes depicted in their stimuli could have helped to drive prediction. The stimuli in Altmann and Kamide’s (1999) study for the most part avoid this problem by using clip art figures that provide very little event information to participants, but the very unreality that lends strength to their manipulation also reduces the ecological validity of their study and likely increases reliance on verb-related information. When verb and event knowledge are further untangled, as in the present study, stronger or weaker verb constraints have no impact on prediction. Instead, prediction was driven by the combination of the scene and the verb regardless of the verb’s constraints.

Although the current study found that verb information did not drive prediction differently from event information, this does not mean that comprehenders do not use verb constraints when generating predictions—just that, given the tight relationship between verb and event knowledge, it is extremely difficult to fully separate effects of these two types of information. It seems unreasonable to suggest that, given a strongly-constraining source of information, participants would not take advantage of it to make predictions. Rather, the present results suggest that the presence of constraining verb information doesn't elicit more or faster predictions, regardless of whether or not the information is used. The connection between verb and event knowledge that makes their dissociation impossible in the current study also means that fully untangling their effects is likely impossible.

Previous research shows that SRVs have an earlier effect on language comprehension than equally severe violations of world knowledge (Warren & McConnell, 2007); however, it is impossible to tell whether these early effects of SRVs are found because comprehenders use selectional restriction information predictively. Thus far studies comparing effects of SRVs to effects of world knowledge violations have all been studies of reading, which are unable to distinguish between prediction and integration (Paczynski & Kuperberg, 2012; Warren & McConnell, 2007; Warren et al, in prep), and therefore cannot resolve the exact mechanism that is affected by a verb's semantic requirements. The present results suggest that selectional restrictions do not lead to greater or more predictions than do world knowledge, indicating that the effects of selectional restrictions seen in previous research may be more integrative than predictive.

Predictive processing has also been hypothesized to be closely related to language production (Pickering & Garrod, 2007). Studies investigating the relationship between prediction and production find that, although older adults show little or no predictive behavior (Federmeier & Kutas, 2005; DeLong et al, 2012), a subset of older adults with high verbal fluency scores tend to behave more like college-age participants during prediction tasks (Federmeier, 2007; DeLong et al, 2012). To investigate the relationship between prediction and production, we examined verbal fluency, an individual difference measure that indexes production capabilities, in younger adults. Despite observing a wide range of fluency scores, we found only a tenuous relationship between verbal fluency and prediction. Verbal fluency score predicted proportions of fixations to the target in the two constrained conditions during the first post-noun bin (200-700 ms post noun offset), with an increase in verbal fluency associated with a decrease in fixation proportion. Fixation proportions to the target during this bin were also significantly lower in the two constrained conditions than in the two control conditions. It's possible that participants with higher verbal fluency were more easily able to move on from fixating the target and direct their attention elsewhere, although because no effects of verbal fluency are seen before noun onset it is again impossible to say whether these effects are because participants with higher verbal fluency are making more predictions or are faster at integration. Verbal fluency was also related to speed to fixate the target across all conditions: participants with higher verbal fluency fixated the target more quickly. When considered with the other results of this study, this finding suggests that verbal fluency is related less to prediction and more to overall speed of language access. If participants with higher verbal fluency are more able to access connections between words in the mental lexicon, they might more quickly activate a verb's common arguments, congruent with McRae and colleagues' findings of common participants and location of events

being primed by the verbs that describe them (Ferretti et al, 2001; McRae & Matsuki, 2009). The current study is not able to provide much evidence for the relationship between prediction and production. However, most of the research showing such a relationship has used older adults. It's possible that the relationship between prediction and production, at least measured with verbal fluency, is more easily measured in a cognitively declining population than in younger adults, and thus such a relationship would not be observable in a sample of only younger adults.

We used the visual world paradigm to examine prediction driven by different sources of information, finding that, at least in this study, selectional restrictions do not appear to grant an advantage to prediction over that granted by knowledge about events, a possibility that would be congruent with previous research on violation detection (Warren & McConnell, 2007) and semantic relatedness (Paczynski & Kuperberg, 2012) showing effects of selectional restrictions over world knowledge during language comprehension. Rather, the primary motivator of predictive eye movements in the present study appears to be non-linguistic knowledge about events, consistent with research finding that event knowledge is used very early during language comprehension (McRae & Matsuki, 2009). Although verbal fluency, a measure of production, was related to latency to fixate the target, there was very little evidence in this study for a relationship between prediction and production. Although future research may make steps towards untangling the contributions of verb and event knowledge to language prediction, the two sources of information are so tightly intertwined that a full dissociation is likely impossible.

## BIBLIOGRAPHY

- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- Bogdanoff, Michael G. and Katz, Richard C. *Modification of the Cloze Procedure for Measuring Reading Levels in Aphasic Adults*. In Clinical Aphasiology Conference: Clinical Aphasiology Conference (1983 : 13th : Phoenix, AZ : May 29-June 2, 1983) / : BRK Publishers(1983), pages 28-34.
- Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, 112(4), 417-436.
- DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language*, 121(3), 226-239.
- Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, 42(2), 133-141.
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491-505.
- Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(1), 182-196.
- Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, 44(4), 516-547.
- Hintz, F., Meyer, A. S., & Huettig, F. (2014). The influence of verb-specific featural restrictions, word association, and production-based mechanisms on language mediated anticipatory eye movements. In *CUNY 2014 Papers & Poster Abstract Booklet*. Paper presented at The CUNY Conference on Human Sentence Processing, The Ohio State University, Columbus, 13-15 March December (pp. 31).

- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137(2), 151–171.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. (1st ed., pp. 275-293). New York, NY: Oxford University Press.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156.
- Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 30(2), 170-210.
- Kukona, A., Fang, S., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, 119(1), 23–42.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, 22, 1-75.
- McRae, K., & Matsuki, K. (2009). People use their knowledge of common events to understand language, and do so as quickly as possible. *Language and Linguistics Compass*, 3(6), 1417–1429.
- McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition*, 33(7), 1174-1184.
- Meyers, E. B., & Blumstein, S. E. (2005). Selectional restriction and semantic priming effects in normals and broca's aphasics. *Journal of Neurolinguistics*, 18(3), 277–296.
- Nakano, H., & Blumstein, S. E. (2004). Deficits in thematic integration processes in broca's and wernicke's aphasia. *Brain and Language*, 88(1), 96–107.
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension?. *Trends in Cognitive Science*, 11(3), 105–110.
- Spivey, M. J., Tanenhaus, M. J., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, 45, 447-481.
- Staub, A., Abbott, M., & Bogartz, R. S. (2012). Linguistically guided anticipatory eye movements in scene viewing. *Visual Cognition*, 20(8), 922-946.
- Stick, S. L., & Hartwig, M. (1977). Use of the cloze procedure to determine morphological or semantic based treatment for aphasics. *Human Communication Research*, 2(3), 177-185.

- Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review*, 14(4), 770-775.
- Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 34(4), 1001-1010.
- Warren, T., Milburn, E., Patson, N. D., & Dickey, M. W. (under review). Comprehending the impossible: What is the role of selectional restriction violations?
- Wlotko, E. W., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology*, 49(6), 770–785.